© Copyright Michael Stubbs 2005.

This pre-publication version 26 October 2005.

The final version was published in

A Gerbig & A Müller-Wood eds (2006) *How Globalization Affects the Teaching of English*. Lampeter: Mellen. 15-29.

It is a revised version of a paper read to the conference on English teaching, University of Trier, Germany, 2 - 4 October 2003.

ON TEACHING CRITICAL RATIONALISM: RECONCILING LINGUISTIC AND LITERARY TEXT ANALYSIS

Michael Stubbs

The introduction to this book talks of a "crisis" in English Studies, and of the need to reconcile literary, linguistic and cultural studies. I will begin with a critical view of some of these buzz words, and will then make some concrete proposals about what might be reconciled with what.

First, the debate is not new, and indeed English Studies seems to suffer crises of confidence around once per generation. Twenty years ago there was an influential book with a similar title to this one: *Re-Reading English* (Widdowson 1982), which started with the editor's essay on "the crisis in English Studies". (This referred to the bitter debate in Cambridge at the time over what should constitute an English syllabus.) In the letters pages of *The London Review of Books*, that discussion sparked what has been called "the most famous shoot-out since the OK Corral". But then Widdowson tells the anecdote, against himself, that he was one day routinely demolishing the literary canon in class, when a student burst out that it was all right for him: "You've had the opportunity to read these books, I haven't, and that's what I've come to university to do". (These quotes are from Day 2003.)

Second, the debate about cultural aspects of English Studies is also not new. In its modern form, it was launched between the 1950s and the 1970s by Richard Hoggart, Raymond Williams and Stuart Hall, partly as a reaction against a previous tradition from the 1930s to the 1950s in which F R Leavis tried to relate textual and social criticism. (Williams 1961 gives an influential account.) One can trace these debates further back still, to the 1870s and to Matthew Arnold on *Culture and Anarchy*, with its arguments about mass culture and the place of English teaching in a moral and cultural education. Or one can trace them forward via the moral panic and the enormous press coverage of English teaching in

schools in the UK, which was triggered by the Kingman and Cox Reports in the late 1980s. (Stubbs 1989 gives an account from one point of view.) These debates have continued under both Conservative and Labour governments. One further squabble started in April 2003 when the (Labour) Secretary of State for Education, Charles Clarke, questioned the value of medieval history, and criticized academics who think the university is a "community of scholars unfettered by the problems of wider society". The value of historical knowledge was quickly pointed out to him, but the debate continues.

Relevance = transferable skills?

Third, worries about a crisis in English Studies amount to worries about its "relevance", and this in turn is part of the centuries-old debate about how to balance a general education and a vocational training. One current interpretation of "relevance" is "transferable skills", which is one more formulation of the opposition between vocational skills and educational knowledge. I will argue that it is a rather superficial formulation, which should be strongly resisted in favour of a much deeper concept of relevance which English Studies has to offer.

Indeed, I am puzzled as to why some colleagues seem to have lost confidence in the value of their own subject, and think they have to try to prepare students for the labour market, rather than teach what they are professionally trained in: namely, English Studies. Not only can too much emphasis on skills and training, rather than education, lead to a very conservative curriculum. In addition, there is evidence that commerce and industry would prefer universities to do what they are good at, and leave job training till later: universities cannot teach the skills required in the workplace, since they do not know what is required.

One specific danger is that a concept of transferable skills can lead to a demand for a type of "computer literacy" which can also be interpreted very superficially. Certainly all students of English should be very familiar with what computers can and cannot do, and I will discuss this below. But this means much more than using computers to prepare PowerPoint presentations. If such views are combined with increasingly frequent calls for project work, then problems can arise. I looked at one student project recently which originated in a seminar on early modern literature. It was a CD-ROM presentation of material on early modern London, largely maps of the city and other visual material. It was visually very impressive, it demonstrated skill in using web-page software for graphic design, and it presented interesting background historical and cultural material. But it had little literary content, in fact very little linguistic text, hardly any connected text, and no argumentative text at all. In a word, the project had almost entirely substituted visual presentation for literary and linguistic content. Worse: when I talked to the teacher on the course about this project, she pointed out that the work had been done by a very good student, who had been forced into this mode of presentation, which privileged style over content, precisely by the pressure to produce a "project" in a particular form. Not all aspects of the English curriculum are equally appropriate for project work. And a demand for projects

means there is less time for the central task of students of English: interpreting and producing texts.

Only connect ...

Although debates about the English curriculum keep returning in only slightly different forms, the language-literature divide has never been resolved. Many students see no connection between the two, and although the longer term aims of university English Studies may seem very obvious to us as teachers, they remain obscure to students. From a student point of view, the curriculum often lacks coherence.

There are some rather obvious mid-way points between lang and lit, such as the linguistic stylistics of literary texts, but then stylistics lies in an academic no-man's land. Within linguistics, there was optimism in the 1960s and 1970s, in the work of Roman Jakobson and Michael Halliday, but stylistics has never made it into the mainstream, since linguistics has never developed a convincing textual theory. Within literary study, there has always been scepticism of linguistic description, which was famously attacked by Stanley Fish. He argued that it is inherently circular: stylisticians pick out linguistic features of texts which they know are important, describe them, and then claim they are important. This criticism has never been answered entirely convincingly by linguists. (One simple answer is attempted by Stubbs 2005.)

A modest proposal

Here is a modest proposal (which I mean to be taken more seriously than Swift's modest proposal). The ideas are so modest, that they may seem very obvious and banal. In a way I hope they do, because that would mean that you agree with them, and it might even mean that they are true. I will here be following the maxim that "some truisms are true". However, even if they seem like truisms to you, I am prepared to bet that the average student could not make the connections explicit.

My main argument is that recent developments in both literary and linguistic theory provide substantial common ground between literary and linguistic studies in a training in textual analysis. We all work with texts: in language teaching, in linguistics teaching and in literature teaching. We are concerned with textual units of different sizes: from phonemes and morphemes, via words, phrases and sentences, to individual texts (such as newspaper articles or novels), and beyond that to text-types (such as expository and narrative), and to intertext. There are many different perspectives, we probably all have our own views about which are more valid or interesting than others, and we differ in how systematic and objective we think such analysis can be. I will ignore all that here, and just assume that there is no single perspective which is uniquely valid, but that we are all interested in texts and their interpretation.

There is a particular German-language tradition in hermeneutics. One could start here from Martin Luther and his demand that people read the Bible for themselves - *sola scriptura* - and not accept interpretations passed down from the established Church. Or one could mention all the work from the early 1800s onwards, from Friedrich Schleiermacher, via Wilhelm von Humboldt and Wilhelm Dilthey, to Hans-Georg Gadamer and Jürgen Habermas. Certainly German students should know something about this, though it would have to treated carefully, otherwise it becomes just another topic on the syllabus, unrelated to anything else.

More generally, students should be explicitly introduced to the difficult set of concepts around text and intertext, translation and interpretation, prescription and description, subjectivity, inter-subjectivity and objectivity, qualitative and quantitative evidence, theory, corroboration and counter-example, and different forms of deductive and inductive argument. Now that linguistics has moved definitively away from its recent mentalistic phase, literary and linguistic studies have a major opportunity to collaborate on empirical documentary methods. This is the main optimistic point in my argument.

Mind the gap ...

Here are some more concrete comments. The title of the book tells us that we should be reconciling things[1]. What is it that we want to reconcile with what? The briefest summary I can manage is that we want to reconcile text and context, but there are at least two huge gaps between text and context. First, two things which have never been reconciled are

- a close attention to the linguistic features of an individual text
- an understanding of the historical and cultural background of the text.

Second, much stylistic analysis tackles only small fragments of text: short poems and short fragments of longer texts. Stylisticians often take a famous, but very short, fragment out of a long book, and therefore out of context. A favourite is the famous opening two pages of Charles Dickens' *Bleak House* (a symbolic description of London in the fog): they analyse these pages in detail, but then ignore the remaining 800 pages and 350,000 words.

A criterion for any analysis of anything must be comprehensive coverage of the data (at a given level of description), but as Phillips (1989: 8) points out, the historical impossibility of studying large amounts of text meant that linguistics was often restricted to studying "the extent of language that can comfortably be accommodated on the average blackboard". For a long time, it was difficult to reconcile the demands of cultural, literary and linguistic theory, because much linguistics, from the 1960s to the 1990s, was restricted to a Chomskyan mentalist mode, and to a narrow concept of language based on a blatant refusal to discuss

[1] Note added 2016. Actually it doesn’t: the title of the book was changed after I had submitted my article ...

authentic texts and corpora. This refusal was, in turn, based on faulty arguments that these represent mere performance. This was a very odd and crude misunderstanding, and due to a simple logical error: a corpus is not performance, but a record of performance (which has been designed, according to a theory of language variation, for linguistic study). Chomskyan approaches have also been based on a concept of linguistic creativity, whereas it is obvious from the analysis of authentic text that a large amount of language use (also in literary works) is repetitious and routine. Very simply, if something is understood, then this is because it is a repetition of something. In this sense, all language use is intertextual (and therefore all corpus study is diachronic).

However, recent developments in empirical linguistics mean that there is now a genuine text- and corpus-based alternative to a Chomskyan approach. This means in turn that the concept of intertextuality, which has been primarily developed within literary studies, can now be studied in empirical and quantitative ways using computer-assisted corpus methods. It took linguistics around a hundred years to progress from studying little things such as phonemes, to studying big things such as long texts and text collections. The largest textual objects which linguists currently look at are corpora: that is, text collections, usually of millions - and often of hundreds of millions - of running words. (There is currently - in 2005 - intense debate over the question of how the billions of pages of texts in the World Wide Web can be systematically selected and exploited for language research: Fletcher in press.) There are detailed discussions elsewhere of precisely how corpora are designed and selected, and whether they can be balanced or representative samples of a language. I will just say that it is possible to have text collections which are samples of general English, of many text-types (such as contemporary fiction, newspaper articles, expository texts, and so on), or of a large sample of nineteenth century authors, or simply of a given author (e.g. a 100 per cent sample of Jane Austen's novels), and so on. Such corpora can then be searched for lexical and grammatical patterns. These new possibilities are often called "corpus linguistics". This is not a very good term, since the aim is not to study corpora: corpora are just samples of data. A better term is empirical or data-intensive linguistics. (Stubbs 2004 reviews these issues.)

Recent work on large corpora has led to significant advances in understanding three aspects of language in use, the relations between

- words and their contexts (more accurately co-texts)
- texts and other texts or text corpora
- routine language use and culture.

A sample project on *Heart of Darkness*

I now have a problem of presentation. If I continue in this general and abstract way, then my argument will be general and abstract. Alternatively, a detailed analysis of a particular long text would take up much more space than I have

available. So what I will do is take an example of a possible student project which arose during a course on textual theory, which I taught with a literary colleague. The seminar had studied Joseph Conrad's *Heart of Darkness*, along with different linguistic and literary analyses of the book (including Stubbs 2005) and articles about textual theory. The aim of the project is to combine

- textual theory, both literary and linguistic
- the central literary and linguistic understanding of a specific / unique text
- general explanatory, presentational and design skills
- specific skills involved in working with hypertext.

It was Saussure who pointed out that words do not refer directly to things in the world, but contract relations with other words in the language system. Similarly, literary texts acquire meaning from their relations with other texts, genres, literary traditions and conventions, and general language use. Allen (2000) provides a good overview, and discusses very clearly how linguistic work by Saussure led to literary and cultural theories of intertextuality.

As a practical and concrete illustration of intertextuality, I suggested to students that they should download the text of Joseph Conrad's *Heart of Darkness* from the web (e.g. from Project Gutenberg), and prepare a hypertext version which shows some of its intra- and intertextual references. If you think (with Barthes 1977) that a text is "a tissue of quotations drawn from the innumerable centres of culture", then such a hypertext is a concrete demonstration that a text is open, and that "the claim to decipher a text becomes quite futile" (p.146-47). (I suppose Barthes means "definitively decipher".)

Examples of categories of links which could be built in include:

1. INTRATEXTUAL REFERENCES: i.e. links between different parts of the text, which a skilled reader could be expected to recognize. Such intratextual patterns could be documented by links to concordance lines which show the collocations and connotations of words in the text.

2. INTERTEXTUAL REFERENCES: e.g. (1) Links to other stories by Conrad, e.g. those which have Marlow as narrator. (2) Links to the Biblical source of specific phrases or general vocabulary and phraseology. (3) Links to phrases in other texts, including novels by Charles Dickens and Jules Verne.

3. GENRE REFERENCES: e.g. links to characteristic texts in a given genre (e.g. Gothic horror) which use similar vocabulary or have similar figures (such as the fool or harlequin).

4. REFERENCES TO CHARACTERISTIC PLOTS: e.g. links to stories by Edgar Allan Poe which turn around the fascination of the abyss.

5. ROUTINE PHRASEOLOGY WHICH ECHOES GENERAL LANGUAGE USE: e.g. a link to a concordance of *over the edge* (a key phrase in the book) which shows its collocates and connotations of danger.

6. CRITICISM: e.g. links to critical discussions of *Heart of Darkness*, such as Chinua Achebe's famous discussion of the racist stereotypes in Conrad's book. (Of course "stereotype" is itself an intertextual concept.)

The essential point about such a project is that it requires presentational skills (including web design, HTML, and so on). However, students have to use these practical skills to illustrate the kind of theoretical knowledge which is their main concern as students of linguistics and literature:

- literary genres, related literary texts, critical secondary literature, etc
- open-ended texts, citations, allusions, inter-text, etc
- corpus linguistic methods: concordances, reference corpora, etc.

My paper to the students also warned them of copyright problems in such projects. And it gave them a reference to Allen (2000: 202-03) which describes a project on a hypertext version of Tennyson's *In Memoriam* by George P Landow (http://www.eastgate.com/catalog/InMemoriam.html).

I think we often assume that students are sophisticated about the use of PCs, of the world-wide-web, and so on. But this is not so. Most students of the humanities use PCs as fancy typewriters or presentation devices (e.g. for PowerPoint). They generally know nothing about the nature of HMTL as a system for describing textual structure (often because they use web-page software which hides the underlying system from the user), or about what kinds of text analysis, some of it quantitative and statistical, are possible with software.

I have chosen the example of *Heart of Darkness* in order to describe a project which combines linguistic and literary analysis. I could have described other more specialized student projects in corpus linguistics, which involve the use (or writing) of text analysis software, and then the design, collection and analysis of substantial corpus data. Two projects of this kind started as term papers and led to publications on extracting specialist terminology from academic articles and on using phraseology to discriminate between text-types (Pusch & Stubbs 2003, Stubbs & Barth 2003).

On evaluating analyses

My experience is also that students of English Studies are very unsure about how to evaluate textual analyses. They are often working within an approach which has no tradition of such systematic criticism and no general criteria for such evaluations. So, my final comments concern how an empirical approach to texts fits into a more general critical rationalism. The kind of textual analysis which I

have proposed has to be critically evaluated by contrasting it with alternative positions such as the following.

- Scepticism / rejection. A stylistic analysis of a literary text adds nothing to a close reading. Worse, it is circular, since it picks out a few linguistic features (which we knew were important), describes them, and then tells us they are important. (This is the Fish critique.)
- A weak defence. A stylistic analysis tells us nothing that literary critics do not already know, but it allows us to describe things more precisely and systematically.
- A stronger defence. A stylistic analysis tells us something new: it can discover features of a literary text which neither average readers nor expert literary critics have noticed.
- The strongest(?) defence. A systematic stylistic analysis can not only describe new things, but can also help to explain readers' reactions to the text (for example, by relating these reactions to unconscious linguistic knowledge: about norms of language use, stylistic variation, routine phraseology, etc).

In addition, we need general criteria for evaluating computer-assisted analyses in the humanities. If it is worth while doing such analysis, then (as Kenny 1992 argues), this is because:

- it is impossible to do it without a computer
- it can make a scholarly contribution in the relevant field (here, literary criticism).

There is no point in using a computer to analyse a literary text, if the analysis can be done by hand and/or if it produces results which are of no literary interest. We also need comparable evaluation criteria for literary and/or cultural approaches to texts. If it is not possible to state evaluation criteria for an approach, then this is a serious defect in the approach.

Critical rationalism

Here are some concluding comments. I have no problem with the notion that university education should be "relevant", but I have severe problems with the more short-term vocational variants which are often proposed. Literature courses have always taught about textual and cultural interpretation: this is always relevant. And linguistics courses have always taught about different kinds of evidence and argument: this is equally relevant. I am frequently astonished at how little my students know about some very basic ideas in the philosophy of science. The number of times students provide one rather shaky piece of data in an essay, and then tell me they have proved something, would make your hair stand on end. A training in critical rationalism (Popper 1994) is always relevant. "Relevance" means a deep study of

- different types of argument, inductive and deductive
- the concept of counter-example
- different types of evidence, qualitative and quantitative
- tools of analysis, including different software tools
- evaluation criteria (which must be applied to both linguistic and literary analysis)
- ... a critical approach in general.

Some elementary statistical concepts (e.g. samples, frequencies, probabilities, norms, simple ways of representing - and misrepresenting - quantitative data) do not require any mathematical ability, just a bit of clear thinking, and should be part of the intellectual equipment of any educated person.

Students of language and literature are trained how to analyse masses of complex textual material which have no single definitive interpretation. This was pointed out a long time ago by Chomsky (1967), when he was starting to analyse the masses of documentary material on the Vietnam War.

In conclusion: I have made a very modest proposal for a syllabus which is based around the concept of text analysis. Our central responsibility in English Studies is to teach language, literature and linguistics: to make explicit the kind of evidence and methods we work with, and to point to links which are probably self-evident to us, but are often completely obscure to students. And I have given one brief illustration of a project which demands not only linguistic and literary analysis, but also the presentational skills which are only possible with modern computer technology.

Reconcile is one of a set of verbs which express similarities, connections or differences between things (Francis et al 1996: 61). (And if you wish, I can provide corpus evidence that this is its normal use.) The word *reconcile* is used when there are difficulties caused by a mismatch between two or more apparently contradictory or conflicting demands which have to be reconciled with or to each other. Ideally one should be able to show that they are not really contradictory.

One main unsolved problem of text analysis is how a close attention to the structure of the text can be reconciled with a full understanding both of normal language use, and also of the cultural and historical background of the text and the reader. Linguists are quite right to insist on the systematic description of the text. Literary critics are quite right to insist on the context and the reader. I have argued that new forms of evidence from the computer-assisted analysis of large corpora can help to fill the yawning gap between these approaches.

Some teachers of English Studies seem to have lost confidence in their central responsibilities, and have been diverted into providing external, vocational justifications for English Studies, and therefore into finding alternatives to English Studies. What they should be doing is finding arguments for the centrality of their

discipline, and this lies in the central importance of understanding difficult texts. It should not be forgotten that both the Thatcher and the Blair governments in the UK have tried, very explicitly, to severely restrict the study of language, history and culture.

When I am in a pessimistic mood, I think that all the current talk of reformed, Europe-wide BAs and MAs will merely lead to courses which leave language, linguistics and literature as separate as they have always been and - worse - which have superficial, bolted-on vocational units. When I am in a more optimistic mood, I think that we have new ways of studying text and intertext, and that we are in the best position for decades

- to relate some central concerns of language, linguistics and literature teaching
- to demonstrate the "relevance" of these central concerns (or, if you like, the "transferable skills" which have always been part of such an education).

This does not require new and different courses, but it does require that we all make explicit to students two things:

- language, linguistics and literature all involve text analysis of different kinds
- textual analyses must be open to evaluation by a range of clearly formulated criteria.

Only in these ways can we make explicit to students the relations between our courses.

References

Allen, G. (2000) *Intertextuality*. London: Routledge.
Barthes, R. (1977) *Image Music Text*. London: Fontana.
Chomsky, N. (1967) The responsibility of intellectuals. *New York Review of Books*, 23 Feb 1967 (and republished in several other places).
Day, G. (2003) Column. *Times Higher Education Supplement*, 3 Jan 2003: 13.
Fish, S. (1973) What is stylistics and why are they saying such terrible things about it? In S. Chatman ed (1973) *Approaches to Poetics*. Columbia UP. Also in D C Freeman ed (1981) *Essays in Modern Stylistics*. London: Methuen.
Fletcher, W (in press) Facilitating the compilation and dissemination of ad-hoc web corpora. In G. Aston, S. Bernardini & D. Stewart eds *Papers from the Fifth International Conference on teaching and Language Corpora*. Amsterdam: Benjamins.
Francis, G. et al (1996) *Grammar Patterns 1:Verbs*. London: HarperCollins.
Kenny, A. (1992) Computers and the humanities. Ninth British Library Research Lecture. The British Library.

Phillips, M. (1989) *Lexical Structure of Text*. Discourse Analysis Monograph 12, English Language Research, University of Birmingham.

Popper, K. R. (1994) *The Myth of the Framework*. London: Routledge.

Pusch, A. & Stubbs, M. (2003) Frequent terms in linguistics: a pilot corpus study for a pedagogical word-list. In C. Tschichold ed *English Core Linguistics*. Bern: Peter Lang. 247-68.

Stubbs, M. (1989) The state of English in the English state. *Language and Education*, 3, 4: 235-50.

Stubbs, M. (2004) Language corpora. In A. Davies & C. Elder eds *Handbook of Applied Linguistics*. Oxford: Blackwell. 106-32

Stubbs, M. (2005) Conrad in the computer: examples of quantitative stylistic methods. *Language and Literature*, 14, 1: 5-24.

Stubbs, M. & Barth, I. (2003) Using recurrent phrases as text-type discriminators: a quantitative method and some findings. *Functions of Language*, 10, 1: 65-108.

Widdowson, P. (1982) The crisis in English studies. In P. Widdowson ed *Re-Reading English*. London: Fontana.

Williams, R. (1961) *The Long Revolution*. London: Chatto & Windus.